Pergamon

Appl. Math. Lett. Vol. 9, No. 2, pp. 91–95, 1996

0893-9659/96 \$15.00 + 0.00

S0893-9659(96)00019-5

A Fermat Principle on Lorentzian Manifolds and Applications

F. ANTONACCI AND P. PICCIONE
Fac. Ingegneria, Dip. di Energetica
Universitá dell'Aquila
Localitá Monteluco, 67040 Roio Poggio (AQ), Italy
flavia@dsiaq1.ing.univaq.it, piccione@dsiaq1.ing.univaq.it

(Received and accepted June 1995)

Abstract—We present a version of the Fermat Principle to Lorentzian manifolds endowed with a time function. The principle is used to obtain some results concerning the existence and multiplicity of light rays, generalizing part of the work in [1–3]. At this time, the results are announced and discussed, while the details of their proofs are left to a forthcoming paper [4].

Keywords—Fermat's Principle, General Relativity, Light rays.

1. INTRODUCTION AND STATEMENT OF THE RESULTS

The Fermat Principle in classical Optics states the trajectory of a light ray joining two points inside a viscous medium is the one for which the *time of percurrence* is minimal. Extensions of the Fermat Principle in General Relativity have been widely investigated by many authors; we would like to mention here the works of Perlick [5,6] which were the starting point of our study. For a more complete reference, the reader can look at [1,2,7] and the references therein.

In the mathematical framework of General Relativity, the trajectories of light rays are given by light-like geodesics on a Lorentzian manifold $\mathcal{M}$, whose points are the *events* of the space-time. Roughly speaking, giving a Fermat Principle means choosing an appropriate manifold $\mathcal{L}_{p,\gamma}$ consisting of paths in $\mathcal{M}$ joining the event p (the source of light) with the time-like curve γ (the observer), and a smooth functional F on $\mathcal{L}_{p,\gamma}$ whose critical points correspond to light-like geodesics.

A Lorentzian manifold $\mathcal{M}$ is said to be *time-oriented* if there exists a time-like smooth vector field Y on $\mathcal{M}$. If $\mathcal{M}$ is time-oriented, a smooth curve $z\colon [0,1] \longmapsto \mathcal{M}$ is said to be *future pointing* (past pointing) if $\langle \dot z(s), Y(z(s))\rangle$ is negative (positive) for all s. If γ is a time-like curve and p is an event in $\mathcal{M}$, we denote by $\mathcal{L}^+_{p,\gamma}$ ($\mathcal{L}^-_{p,\gamma}$) the set of all C^2, future pointing (past pointing), light-like curves $z\colon[0,1] \longmapsto \mathcal{M}$ joining p and γ:

$$\mathcal{L}^+_{p,\gamma} = \left\{z\colon[0,1]\longmapsto \mathcal{M} \mid z \in C^2\left([0,1],\mathcal{M}\right),\ \langle \dot z, \dot z\rangle \equiv 0, \langle \dot z, Y(z)\rangle < 0, \right.$$
$$\left. z(0) = p,\ z(1) \in \operatorname{supp}(\gamma)\right\}.$$

REMARK. Notice that the space $\mathcal{L}^+_{p,\gamma}$ may be empty. For instance, if $\mathcal{M} = \mathbb{R}\times\mathbb{R}$ is equipped with the Lorentzian metric

$$\mathrm{d}s^2 = \left(1+t^2\right)^2 \mathrm{d}x^2 - \mathrm{d}t^2,$$

then there exists no light-like curve joining $p = (0,0)$ with the time-like curve $\gamma(s) = (\frac{\pi}{2}, s)$.

Typeset by $\mathcal{A}_{\mathcal{M}}\mathcal{S}$-TeX

We will assume throughout the paper that p and γ are chosen so that $\mathcal{L}^+_{p,\gamma}$ ($\mathcal{L}^-_{p,\gamma}$) is not empty.

We introduce the following functionals on $\mathcal{L}^+_{p,\gamma}$:

$$F(z) = -\int_0^1 \langle \dot z(s), Y(z(s))\rangle \, \mathrm{d}s, \tag{1.1}$$

and

$$Q(z) = \int_0^1 \langle \dot z(s), Y(z(s))\rangle^2 \, \mathrm{d}s. \tag{1.2}$$

A *time function* on $\mathcal{M}$ is a smooth function $T\colon \mathcal{M} \longmapsto \mathbb{R}$ whose (Lorentzian) gradient ∇T is time-like. If $\mathcal{M}$ has a universal time function T, then $\mathcal{M}$ is time-oriented by the vector field $Y = -\nabla T$. Observe that, if Y is given by the gradient of the time function T, then the functional F is given by

$$F(z) = \int_0^1 \langle \dot z(s), \nabla T(z(s))\rangle \, \mathrm{d}s = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}s}\,(T(z(s))\, \mathrm{d}s = T(z(1)) - T(z(0)). \tag{1.3}$$

Observe that both F and Q are nonnegative functionals on $\mathcal{L}^+_{p,\gamma}$, and thus they are bounded from below.

Recall that a smooth curve z in $\mathcal{M}$ is called a pregeodesic if there exists a reparametrization of z which is a geodesic in $\mathcal{M}$, or, equivalently, if $\nabla_s \dot z$ is parallel to $\dot z$, where ∇_s is the covariant derivative along z with respect to the Levi-Civita connection of $\mathcal{M}$.

THEOREM 1.1. *Let $\mathcal{M}$ be a Lorentzian manifold with a time function T, F be the functional given by (1.3) and z be in $\mathcal{L}^+_{p,\gamma}$. Then, z is a critical point of F if and only if z is a pregeodesic.*

An analogous result holds for past pointing light-like pregeodesics in $\mathcal{L}^-_{p,\gamma}$.

If $\mathcal{M}$ is a manifold with a time function T, then any time-like curve can be parametrized by t, so that the Fermat Principle given by Theorem 1.1 seems somewhat less general than Perlick's results in [5,6]. Despite this fact, we would like to point out that in our case, the functional F has a concrete analytic definition, which is easier to handle with standard techniques of Nonlinear Analysis.

THEOREM 1.2. *Let $\mathcal{M}$ be a Lorentzian manifold with a time function T, and let Q be the functional given by (1.2), where $Y = -\nabla T$. A curve $z \in \mathcal{L}^+_{p,\gamma}$ is a critical point of Q if and only if z is a pregeodesic such that $\langle \dot z(s), Y(z(s))\rangle$ is constant.*

An analogous result holds for past pointing pregeodesics.

Light rays pointing in the past of an event are meaningful in General Relativity. See [8] for a reference on an application of the theory to phenomena of *gravitational lensing*.

REMARK. The functionals F and Q resemble, respectively, the length functional L and the energy functional E on a Riemannian manifold. Indeed, if $\mathcal{M}_0$ is a Riemannian manifold and $x_0, x_1 \in \mathcal{M}_0$, then denoting by $C^2(\mathcal{M}_0, x_0, x_1)$ the space of C^2 curves in $\mathcal{M}_0$ joining x_0 and x_1, the critical points of L are exactly the pregeodesics in $\mathcal{M}_0$ joining x_0 and x_1. Moreover, $x \in C^2(\mathcal{M}_0, x_0, x_1)$ is a critical point of E if and only if x is a pregeodesic in $\mathcal{M}_0$ joining x_0 and x_1, with the property that $\langle \dot x, \dot x\rangle$ is constant. As in the Riemannian case, the reader should observe that Q is a functional which is easier to study than F. Similar results are obtained in [3,7] for Lorentzian manifolds which are isometric to an orthogonal splitting.

The study of the critical points of the functionals F and Q is used to prove some results regarding the existence and multiplicity of light rays on Lorentzian manifolds with a time function. In such manifolds, the intrinsic property that guarantees the existence of light rays joining an event p with a time-like curve γ is a sort of compactedness of the manifold $\mathcal{L}^\pm_{p,\gamma}$, explained in the following definition.

DEFINITION. *Let c be a real number. The manifold $\mathcal{L}^+_{p,\gamma}$ is called c-precompact if for every sequence $\{z_n\}$ in $\mathcal{L}^+_{p,\gamma}$ such that $F(z_n) \le c$, there exists a subsequence $\{z_{n_k}\}$ which is uniformly convergent in $\mathcal{M}$, up to a reparametrization.*

DEFINITION. *If $\mathcal{M}$ is time oriented by the vector field Y, a smooth curve $\gamma\colon (a,b) \longmapsto \mathcal{M}$ is called a time-like vertical curve if γ is a maximal solution of the equation*

$$\dot\gamma = Y(\gamma).$$

It is easy to see that a time-like vertical curve γ is a closed embedding of a real interval in $\mathcal{M}$.

We have the following result regarding the existence of light rays joining an event with a vertical time-like curve.

THEOREM 1.3. *Let $\mathcal{M}$ be a Lorentzian manifold with a time function, $p \in \mathcal{M}$, and γ a time-like vertical curve in $\mathcal{M}$, with $p \notin \operatorname{supp}(\gamma)$. Suppose that $\mathcal{L}^+_{p,\gamma}$ is nonempty, and that there exists a number $c > \inf\{F(z) : z \in \mathcal{L}^+_{p,\gamma}\}$ such that $\mathcal{L}^+_{p,\gamma}$ is c-precompact. Then, there exists a light-like, future pointing geodesic in $\mathcal{M}$ joining p with γ. An analogous result holds for past pointing geodesics.*

The multiplicity of light rays is obtained under a *nontriviality* assumption on the topology of $\mathcal{M}$. For the correct statement of the result we will use the space

$$\hat{\mathcal{L}}_{p,\gamma} = \left\{ z \in C^2([0,1],\mathcal{M}) : \langle \dot z, \dot z\rangle = 0,\ \langle Y(z), \dot z\rangle \le 0,\ z(0) = p,\ z(1) \in \operatorname{supp}(\gamma)\right\},$$

and we denote by $\operatorname{cat}(\hat{\mathcal{L}}_{p,\gamma})$ its Ljusternik-Schnirelman category.

THEOREM 1.4. *If $\hat{\mathcal{L}}^+_{p,\gamma}$ is c-precompact for every $c \in \mathbb{R}$, then there are at least $\operatorname{cat}(\hat{\mathcal{L}}^+_{p,\gamma})$ light-like, future pointing geodesics joining p and γ. In particular, if $\operatorname{cat}(\hat{\mathcal{L}}^+_{p,\gamma}) = +\infty$, then there exists a sequence $\{z_n\}$ of light-like geodesics in $\mathcal{L}^+_{p,\gamma}$ such that*

$$\lim_{n\to\infty} F(z_n) = +\infty.$$

An analogous result holds for past pointing geodesics.

2. OUTLINE OF THE PROOFS

In this section, we give an outline of the proofs of the two versions of the Fermat Principle, and we refer to [4] for the rest of the work.

The proof of Theorem 1.1 is essentially the same as Perlick's proof about the critical points of the Arrival Time functional introduced in [6].

Let $\mathcal{M}$ be a Lorentzian manifold with a time function T, let $Y = -\nabla T$ be its (Lorentzian) gradient, p an event of $\mathcal{M}$, and $\gamma\colon [a,b] \longmapsto \mathcal{M}$ a smooth, time-like curve, with $p \notin \operatorname{supp}(\gamma)$. It is easy to see that $\mathcal{L}^+_{p,\gamma}$ is a smooth submanifold of $C^2([0,1],\mathcal{M})$ (with the same regularity of γ).

For $z \in \mathcal{L}^+_{p,\gamma}$, since $z(1) \in \operatorname{supp}(\gamma)$, then there exists a number s_z such that $z(1) = \gamma(s_z)$. Observe that s_z is just the Arrival Time functional defined by Perlick (see [5,6]).

The tangent space $T_z\mathcal{L}^+_{p,\gamma}$ at the point z can be identified with the space

$$T_z\mathcal{L}^+_{p,\gamma} = \Big\{ \zeta\colon [0,1] \longmapsto T\mathcal{M} \mid z \in C^2([0,1],\mathcal{M}),\ \zeta(s) \in T_{z(s)}\mathcal{M},$$
$$\langle \nabla_s\zeta, \dot z\rangle \equiv 0,\ \zeta(0) = 0,\ \zeta(1) \parallel \dot\gamma(s_z)\Big\}.$$

Let F be the smooth functional on $\mathcal{L}^+_{p,\gamma}$ defined by (1.1). For every $z \in \mathcal{L}^+_{p,\gamma}$ and every $\zeta \in T_z\mathcal{L}^+_{p,\gamma}$, one has $F'(z)[\zeta] = \langle \nabla T(z(1)), \zeta(1)\rangle$, and we have the following characterization of the tangent space to $\mathcal{L}^+_{p,\gamma}$ at a critical point for F.

LEMMA 2.1. *$z \in \mathcal{L}^+_{p,\gamma}$ is a critical point for F if and only if for every $\zeta \in T_z\mathcal{L}^+_{p,\gamma}$, it holds*

$$\langle \nabla T(z(1)), \zeta(1)\rangle = 0.$$

∎

PROOF OF THEOREM 1.1. Let $z \in \mathcal{L}^+_{p,\gamma}$ be a critical point for F, so that, from Lemma 2.1, for every $\zeta \in T_z\mathcal{L}^+_{p,\gamma}$, we have $\langle \nabla T(z(1)), \zeta(1)\rangle = 0$. Let U be a the vector field given by the parallel transport of $\dot\gamma(s_z)$ along z; i.e., U satisfies

$$\nabla_s U = 0,$$
$$U(1) = \dot\gamma(s_z).$$

Since the parallel transport is an isometry, then $U(s)$ is time-like for every s.

For any $W \in C^\infty_0([0,1], T\mathcal{M})$ such that $W(s) \in T_{z(s)}\mathcal{M}$ for all s, the vector field $\zeta_W(s)$ along z given by

$$\zeta_W(s) = W(s) - \left(\int_0^s \frac{\langle \nabla_s W, \dot z\rangle}{\langle \dot z, U\rangle}\,\mathrm{d}r\right) U(s)$$

is in $T_z\mathcal{L}^+_{p,\gamma}$. Clearly, $\zeta_W(0) = 0$ and $\zeta(1) \parallel \dot\gamma(s_z)$. Moreover, a straightforward calculation shows that $\langle \nabla_s\zeta_W, \dot z\rangle \equiv 0$, so that $\zeta_W \in T_z\mathcal{L}^+_{p,\gamma}$. The equality $\langle Y(z(1)), \zeta_W(1)\rangle = 0$ gives

$$\left(\int_0^1 \frac{\langle \nabla_s W, \dot z\rangle}{\langle \dot z, U\rangle}\,\mathrm{d}r\right)\langle U(1), Y(z(1))\rangle = 0.$$

Since $U(1) \parallel Y(z(1))$, then it must be

$$\int_0^1 \frac{\langle \nabla_s W, \dot z\rangle}{\langle \dot z, U\rangle}\,\mathrm{d}r = -\int_0^1 \left\langle W, \frac{\mathrm{d}}{\mathrm{d}s}\frac{\dot z}{\langle U, \dot z\rangle}\right\rangle \mathrm{d}r = 0, \qquad \forall\, W \in C^\infty_0([0,1], T\mathcal{M}),$$

or, equivalently,

$$\frac{\mathrm{d}}{\mathrm{d}s}\frac{\dot z}{\langle U, \dot z\rangle} = \frac{\nabla_s \dot z}{\langle U, \dot z\rangle} + \dot z\,\frac{\mathrm{d}}{\mathrm{d}s}\frac{1}{\langle U, \dot z\rangle} = 0.$$

In particular, $\nabla_s \dot z$ is parallel to $\dot z$ and z is a pregeodesic.

Conversely, let $z \in \mathcal{L}^+_{p,\gamma}$ be a pregeodesic, so that $\nabla_s \dot z = \lambda(s)\dot z$ for some smooth function λ. Take $\zeta \in T_z\mathcal{L}^+_{p,\gamma}$, so that $\langle \nabla_s\zeta, \dot z\rangle \equiv 0$. A direct calculation shows that the function $\phi_\zeta(s) = \langle \zeta(s), \dot z(s)\rangle$ satisfies the Cauchy problem

$$\dot\phi = \lambda\phi,$$
$$\phi(0) = 0,$$

so that $\phi \equiv 0$. In particular, $\phi(1) = \langle \dot z(1), \zeta(1)\rangle = 0$, which implies that $\zeta(1) = 0$, since $\dot z(1)$ is light-like and $\zeta(1)$ is time-like. Then, from Lemma 2.1, we have the thesis. ∎

The following simple result is needed in the proof of Theorem 1.2.

LEMMA 2.2. *Let $z \in \mathcal{L}^+_{p,\gamma}$ be a pregeodesic. Then, there exists a unique reparametrization $w \in \mathcal{L}^+_{p,\gamma}$ of z such that $\langle \dot w, Y(w)\rangle$ is constant.*

PROOF. The parameter is $\sigma(s) = (\int_0^1 \langle \dot z, Y(z)\rangle\,\mathrm{d}r)^{-1}\int_0^s \langle \dot z, Y(z)\rangle\,\mathrm{d}r$. ∎

OUTLINE OF THE PROOF OF THEOREM 1.2. We set $E(s) = \langle \dot z(s), Y(z(s))\rangle$. We need to prove that $E(s)$ is constant if z is a critical point for Q. Namely, since Y is conservative, one has

$$Q'(z)[\zeta] = 2\int_0^1 E(s)\,\frac{\mathrm{d}}{\mathrm{d}s}\langle Y(z(s)), \zeta(s)\rangle,$$

for every $\zeta \in T_z\mathcal{L}^+_{p,\gamma}$; therefore, if E is constant, then from Theorem 1.1 and Lemma 2.2, the thesis would follow. If $z \in \mathcal{L}^+_{p,\gamma}$ is a critical point for Q and $\zeta \in T_z\mathcal{L}^+_{p,\gamma}$, then, since $\zeta(0) = 0$, integration by parts shows that

$$0 = \frac{1}{2}Q'(z)[\zeta] = \langle E(1)Y(z(1), \zeta(1)\rangle - \int_0^1 \langle E'(s)Y(z(s)), \zeta(s)\rangle \, \mathrm{d}s. \tag{2.1}$$

Arguing as in the proof of Theorem 1.1, for every $W \in C_0^\infty([0,1].T\mathcal{M})$, with $W(s) \in T_{z(s)}\mathcal{M}$ for every s, we define a vector field $\zeta_W(s)$ along z, such that $\zeta_W \in T_z\mathcal{L}^+_{p,\gamma}$, and for which, reversing the order of integration and integrating by parts, (2.1) becomes

$$\begin{aligned}
0 &= -\int_0^1 \langle E'(s)Y(z), W(s)\rangle \, \mathrm{d}s + \int_0^1 \left\langle E'(s)Y(z), U(s)\left(\int_0^s \frac{\langle \nabla_s W, \dot z\rangle}{\langle \dot z, U\rangle}\,\mathrm{d}r\right)\right\rangle \mathrm{d}s \\
&\quad - \left[\int_0^1 \left\langle W(s), \nabla_s\left(\frac{\dot z(s)}{\langle U, \dot z\rangle}\right)\right\rangle \mathrm{d}s\right] \langle E(1)Y(z(1)), U(z(1))\rangle \\
&= -\int_0^1 \langle E'(s)Y(z), W(s)\rangle \, \mathrm{d}s + \int_0^1 \frac{\langle \nabla_s W, \dot z\rangle}{\langle U, \dot z\rangle} \cdot \left(\int_s^1 \langle E'(r)Y(z), U(r)\rangle \,\mathrm{d}r\right) \mathrm{d}s \\
&\quad - \left[\int_0^1 \left\langle W(s), \nabla_s\left(\frac{\dot z(s)}{\langle U, \dot z\rangle}\right)\right\rangle \mathrm{d}s\right] \langle E(1)Y(z(1)), U(z(1))\rangle \\
&= -\int_0^1 \langle E'(s)Y(z(s)), W(s)\rangle \, \mathrm{d}s + \int_0^1 \left\langle W(s), \nabla_s\left[\frac{\dot z}{\langle U, \dot z\rangle}\left(\int_s^1 \langle U(s), E'(s)Y(z)\rangle \,\mathrm{d}r\right)\right]\right\rangle \mathrm{d}s \\
&\quad - \left[\int_0^1 \left\langle W(s), \nabla_s\left(\frac{\dot z(s)}{\langle U, \dot z\rangle}\right)\right\rangle \mathrm{d}s\right] \langle E(1)Y(z(1)), U(z(1))\rangle .
\end{aligned}$$

Since W is arbitrary, we pass from a weak to a strong form of the previous equality which multiplied by $\dot z$ gives

$$Q'(z)[\zeta_W] = 0 \quad \Longleftrightarrow \quad E'(s)E(s) = 0.$$

It follows that $\langle Y, \dot z\rangle$ is constant and we are done. ∎

REFERENCES

1. D. Fortunato, F. Giannoni and A. Masiello, A Fermat principle for stationary space-times and applications to light rays, *Jour. of Geometry and Physics* **515**, 1–30 (1994).
2. D. Fortunato and A. Masiello, Fermat principles in General Relativity and the existence of light rays on Lorentzian manifolds (preprint) (1995).
3. F. Giannoni and A. Masiello, On a Fermat principle in General Relativity. A Ljusternik-Schninerlman theory for light rays (preprint) (1995).
4. F. Antonacci and P. Piccione, Existence and multiplicity of light rays in Lorentzian manifolds with a time function (preprint) (1995).
5. V. Perlick, On Fermat's principle in General Relativity: I. The conformally stationary case, *Class. Quantum Grav.* **7**, 1319–1331 (1990).
6. V. Perlick, On Fermat's principle in General Relativity: II. The general case, *Class. Quantum Grav.* **7**, 1849–1867 (1990).
7. K. Uhlembeck, A Morse theory for geodesics on a Lorentz manifold, *Topology* **14**, 69–90 (1975).
8. P. Schneider, J. Ehlers and E. Falco, *Gravitational Lensing*, Springer-Verlag, Berlin, (1992).
9. M.S. Berger, *Nonlinearity and Functional Analysis*, Vol. 74, Academic Press, New York, (1977).
10. B. O'Neal, *Semi-Riemannian Geometry with Applications to Relativity*, Academic Press, New York, (1983).